BIT 3 (1963), 27–43

A SPECIAL STABILITY PROBLEM FOR LINEAR MULTISTEP METHODS*

GERMUND G. DAHLQUIST

Abstract.

The trapezoidal formula has the smallest truncation error among all linear multistep methods with a certain stability property. For this method error bounds are derived which are valid under rather general conditions. In order to make sure that the error remains bounded as $t \to \infty$, even though the product of the Lipschitz constant and the step-size is quite large, one needs not to assume much more than that the integral curve is uniformly asymptotically stable in the sense of Liapunov.

1. Introduction.

The general linear k-step method for the approximate numerical computation of the solution $\boldsymbol{x} = \boldsymbol{x}(t)$ of a system of ordinary differential equations of the first order,

$$d\boldsymbol{x}/dt = \boldsymbol{f}(t,\boldsymbol{x}),\ \boldsymbol{x}(0) = \boldsymbol{x}_0,\ (\boldsymbol{x} \in R^s,\ t \geqq 0)\ , \tag{1.1}$$

is defined by the formula,

$$\alpha_k \boldsymbol{x}_{n+k} + \alpha_{k-1}\boldsymbol{x}_{n+k-1} + \ldots + \alpha_0 \boldsymbol{x}_n = h(\beta_k \boldsymbol{f}_{n+k} + \ldots + \beta_0 \boldsymbol{f}_n)\ , \tag{1.2}$$

The theory of such methods is treated thoroughly in the book of Henrici [6]. We assume that α_i, β_i are real constants, $i = 0,1,2,\ldots k$, $\alpha_k \neq 0$, h is a positive constant called the step-size, $t_m = mh$, $\boldsymbol{f}_m = \boldsymbol{f}(t_m,\boldsymbol{x}_m)$. If the vectors $\boldsymbol{x}_0,\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{k-1}$ are given, then $\boldsymbol{x}_k,\boldsymbol{x}_{k+1},\ldots$ are computed recursively by (1.2). This offers no difficulties, when the method is *explicit*, i.e. when $\beta_k = 0$. When the method is *implicit*, i.e. when $\beta_k \neq 0$, some conditions on h and $\boldsymbol{f}$ are required in order to guarantee the existence and uniqueness of $\boldsymbol{x}_{n+k}$, when $\boldsymbol{x}_{n+k-1},\ldots,\boldsymbol{x}_{n+1},\boldsymbol{x}_n$ are known, cf. Section 3.

In connection with the difference equation (1.2), it is natural to introduce the polynomials

$$\varrho(\zeta) = \sum_{j=0}^{k} \alpha_j \zeta^j,\ \sigma(\zeta) = \sum_{j=0}^{k} \beta_j \zeta^j \tag{1.3}$$

and the operator

* The preparation of this paper was partly sponsored by the Office of Naval Research and the US Army Research Office (Durham). Reproduction in whole or in part is permitted for any purpose of the US Government.

$$L = \varrho(E) - hD\sigma(E)\,, \tag{1.4}$$

where $D = d/dt$, and E is the displacement operator, defined by

$$Ex(t) = x(t+h), \text{ or } Ex_n = x_{n+1}\,. \tag{1.5}$$

It is assumed that $\varrho(\zeta)$ and $\sigma(\zeta)$ have no common divisor. The *order* of a method is the largest integer p such that

$$L\varphi(t) = 0\,,$$

identically for all polynomials $\varphi(t)$ of degree p. By Taylor's theorem with the remainder in integral form it then follows that

$$L\psi(t) \sim -ch^{p+1}\psi^{(p+1)}(t), \qquad (h \to 0) \tag{1.5}$$

for arbitrary functions $\psi(t)$, $\psi(t) \in C^{p+1}$, where $c \neq 0$, and where c and p are independent of the function $\psi(t)$, although they depend on the coefficients of $\varrho(\zeta)$ and $\sigma(\zeta)$. A method is called *consistent*, when $p \geqq 1$. It is easily shown, cf. [6, p. 224], that the condition of consistency is expressed by the relations

$$\varrho(1) = 0,\ \varrho'(1) = \sigma(1)\,. \tag{1.6}$$

It follows that $\sigma(1) \neq 0$, because ϱ and σ would otherwise have a common factor. The quantity c^*,

$$c^* = c/\sigma(1)\,,$$

which is called the *error constant* of a method, is an adequate measure for the comparison of the accuracy of methods with the same p, cf. [6, pp. 223, 238 and 251].

The constants c (or c^*) and p can be determined by a suitable special choice of $\psi(t)$ in (1.5). Take $\psi(t) = e^t$, and put $e^h = \zeta$. Then,

$$\varrho(\zeta) - \sigma(\zeta)\log\zeta \sim -c\cdot(\zeta-1)^{p+1},\ (\zeta \to 1)\,,$$

whence

$$\log\zeta - \varrho(\zeta)/\sigma(\zeta) \sim c^*\cdot(\zeta-1)^{p+1},\ (\zeta \to 1)\,. \tag{1.7}$$

It is known that, for a given k, the polynomials $\varrho(\zeta)$ and $\sigma(\zeta)$ can be determined so that $p = 2k$, and that no larger p is possible. However, it is natural to require that, if h is small, then x_n should be close to $x(t_n)$ in some sense, for all t_n of interest, for any choice of starting vectors, x_i, $(i = 0, 1, \ldots, k-1)$, sufficiently close to $x(t_i)$. Several exact, idealized definitions of this vague requirement have been suggested in the literature, and it has been found that the maximum value of p has to be reduced considerably by such requirements. For instance, no method with $p > k+2$ can possess a certain stability property, cf. [6, pp. 217 and 229], which it is reasonable to require for any extensive numerical integration.

In this paper we shall investigate a different formulation of this requirement.

DEFINITION. *A k-step method is called A-stable, if all solutions of* (1.2) *tend to zero, as $n \to \infty$, when the method is applied with fixed positive h to any differential equation of the form,*

$$dx/dt = qx\,, \tag{1.8}$$

where q is a complex constant with negative real part.

In most applications A-stability is not a necessary property. For certain classes of differential equations, however, it would be desirable to have an A-stable method with a small truncation error. A simplified example is the numerical integration over a long time of a non-homogeneous linear system with constant coefficients,

$$d\boldsymbol{x}/dt = \boldsymbol{Q}\boldsymbol{x}+\boldsymbol{f}(t)\,,$$

where some of the eigenvalues of $\boldsymbol{Q}$ have large modulus but negative real part. The solution is of the form

$$\boldsymbol{x} = \boldsymbol{g}(t)+e^{\boldsymbol{Q}t}\boldsymbol{c}\,,$$

In many problems $\boldsymbol{g}(t)$ has a relatively slow variation, cf. Dahlquist [4]. When the components of $e^{\boldsymbol{Q}t}\boldsymbol{c}$ in the directions of the eigenvectors mentioned have lost their importance in the physical system, one would like to proceed with a step h, determined only by the behaviour of $\boldsymbol{g}(t)$ and independent of the norm of $\boldsymbol{Q}$. Non-linear problems of a similar type are encountered in many fields, such as control engineering or chemical engineering, cf. Hamming [5, p. 218]. Although it may be worthwhile to design special methods for such problems, it is of interest to see what can be achieved within the class of linear multistep methods. In Section 3, it is shown that the most accurate of all A-stable linear multistep methods has a remarkable stability property even in non-linear problems.

The requirement of A-stability is an extreme formulation of the wishes in such situations. The reader may also enjoy the less extreme approach to this class of problems made by Robertson [7], who designs linear multistep methods (with $k=2$, $p=3$), such that all solutions of (1.2) tend to zero in a large portion of the complex plane for the quantity qh, though not in the whole half-plane $\operatorname{Re}(qh)<0$.

2. Some consequences of A-stability.

A preliminary upper bound for the order of an A-stable method was obtained in [4, Theorem 4]. In this Section, the least upper bound will be found. It is equal to 2. We first need a Lemma.

Lemma 2.1. *A k-step method is A-stable, if and only if $\varrho(\zeta)/\sigma(\zeta)$ is regular and has a non-negative real part for $|\zeta| > 1$.*

Proof. When $f(t,x) = qx$, (1.2) becomes a difference equation with constant coefficients, the characteristic equation of which reads

$$\varrho(\zeta) - qh\sigma(\zeta) = 0 . \tag{2.1}$$

A-stability is then equivalent to the proposition:

$$\text{(2.1), and } \operatorname{Re}(qh) < 0 \text{ implies } |\zeta| < 1 .$$

In other words:

$$\text{(2.1), and } |\zeta| \geqq 1 \text{ implies } \operatorname{Re}(qh) \geqq 0 .$$

However, $qh = \varrho(\zeta)/\sigma(\zeta)$, if $\sigma(\zeta) \neq 0$,

If ζ_1 is a zero of $\sigma(\zeta)$, then $\varrho(\zeta_1) \neq 0$, because ϱ and σ are not allowed to have common factors. In the neighbourhood of ζ_1, one has

$$\varrho(\zeta)/\sigma(\zeta) \sim -a(\zeta-\zeta_1)^{-m}, \qquad a \neq 0 ,$$

for some positive integer m. This is clearly inconsistent with $\operatorname{Re}\{\varrho(\zeta)/\sigma(\zeta)\} \geqq 0$ in a whole circle around ζ_1. Hence $\sigma(\zeta_1) \neq 0$, if $|\zeta| > 1$, i.e. $\varrho(\zeta)/\sigma(\zeta)$ is regular for $|\zeta| > 1$. (Simple zeros of $\sigma(\zeta)$ may, however, exist on the boundary, $|\zeta| = 1$.) This proves the Lemma.

A similar argument gives the following result.

Theorem 2.1. *An explicit k-step method cannot be A-stable.*

Proof. $\beta_k = 0$ for explicit methods. Hence, for some integer m, $m \geqq 1$, we have $\sigma(\zeta) \sim a\zeta^{k-m}$, when $\zeta \to \infty$, $a \neq 0$. However, $\varrho(\zeta) \sim \alpha_k\zeta^k$, $\alpha_k \neq 0$. Hence $\varrho(\zeta)/\sigma(\zeta) \sim b\zeta^m$, where $b \neq 0$, $m \geqq 1$. This is clearly inconsistent with a non-negative real part for all ζ outside the unit circle.

On the other hand, by the aid of Lemma 2.1 and (1.6) it is easy to verify that there exist implicit methods, which are both A-stable and consistent, e.g. the trapezoidal rule, which has the generating polynomials

$$\varrho(\zeta) = \zeta - 1, \quad \sigma(\zeta) = \tfrac{1}{2}(\zeta + 1) . \tag{2.2}$$

Another example is

$$\varrho(\zeta) = \zeta - 1, \quad \sigma(\zeta) = \zeta .$$

The following example shows that there exist consistent and A-stable methods for any positive integer k:

$$\varrho(\zeta) = \zeta^k - 1, \quad \sigma(\zeta) = \tfrac{1}{2}k(\zeta^k + 1) .$$

We shall now prove the main result of this Section.

Theorem 2.2. *The order, p, of an A-stable linear multistep method cannot exceed 2. The smallest error constant, $c^* = \frac{1}{12}$, is obtained for the trapezoidal rule, $k=1$, with the generating polynomials* (2.2).

Proof. Introduce a new variable, z, by the transformation

$$z = (\zeta+1)/(\zeta-1), \qquad \zeta = (z+1)/(z-1)\,,$$

and put

$$\left(\frac{z-1}{2}\right)^k \varrho(\zeta(z)) = r(z) = \sum_{j=0}^{k-1} a_j z^j\,,$$

$$\left(\frac{z-1}{2}\right)^k \sigma(\zeta(z)) = s(z) = \sum_{j=0}^{k} b_j z^j\,.$$

(It follows from (1.6) that $a_k = 0$.) The relation (1.7) for the determination of p and c^* may now be written in the form,

$$\log\frac{z+1}{z-1} - \frac{r(z)}{s(z)} \sim c^* \left(\frac{2}{z}\right)^{p+1}, \qquad (z \to \infty)\,.$$

Expanding the logarithm into powers of $1/z$, we obtain, if $p \geqq 2$,

$$r(z)/s(z) = 2z^{-1} + (\tfrac{2}{3} - 8c')z^{-3} + O(z^{-4}), \qquad (z \to \infty)\,, \tag{2.3}$$

where

$$c' = \begin{cases} c^*, & \text{if } p = 2\,, \\ 0, & \text{if } p > 2\,. \end{cases}$$

We shall see that a positive coefficient of z^{-3} is inconsistent with A-stability. Lemma 2.1 may now be written thus:

A k-step method is A-stable, if and only if $r(z)/s(z)$ is regular and has a non-negative real part in the half-plane $\mathrm{Re}(z) > 0$.

Since the statement is independent of the degrees of the polynomials $r(z)$ and $s(z)$, it is natural to apply a general device from the theory of analytic functions. Following a suggestion of Professor P. D. Lax (oral communication), we shall use a variant of Riesz–Herglotz' theorem, cf. [1, p. 152], according to which any analytic function $\varphi(z)$ satisfying the conditions

$$\sup\{|x\varphi(x)| \mid 0 < x < \infty\} < \infty\,,$$

$$\varphi(z) \text{ regular for } \mathrm{Re}(z) > 0\,,$$

$$\mathrm{Re}\,\varphi(z) \geqq 0 \text{ for } \mathrm{Re}(z) > 0\,,$$

can be represented by an integral

$$\varphi(z) = \int_{-\infty}^{\infty} \frac{d\omega(t)}{z-it}, \qquad (\mathrm{Re}(z) > 0)\,, \tag{2.4}$$

where $\omega(t)$ is bounded and non-decreasing. (For rational functions, this theorem may be derived in an elementary way from Poisson's integral.) We can apply this theorem to $\varphi(z)=r(z)/s(z)$, because of (2.3) and the Lemma.

For x positive. $r(x)/s(x)$ is real. Hence, by (2.4),

$$\frac{xr(x)}{s(x)} = \int_{-\infty}^{\infty} \frac{x d\omega(t)}{x-it} = \int_{-\infty}^{\infty} \frac{x^2 d\omega(t)}{x^2+t^2}, \tag{2.5}$$

Since, for given t, $x^2/(x^2+t^2)$ is a non-decreasing function of x, $xr(x)/s(x)$ is also non-decreasing, which is clearly inconsistent with a positive coefficient of z^{-3} in the expansion (2.3). Hence

$$\tfrac{2}{3}-8c' \leqq 0, \text{ i.e. } c^* = c' \geqq \tfrac{1}{12},\ p = 2 .$$

The minimum of c^* is obtained for the trapezoidal rule, $k=1$, because in this case

$$r(z)/s(z) = \varrho(\zeta)/\sigma(\zeta) = 2(\zeta-1)/(\zeta+1) = 2/z .$$

Hence, by (2.3), $\frac{2}{3}-8c'=0$, i.e. $c^*=c'=\frac{1}{12}$. This completes the proof.

In fact, *the trapezoidal rule is the only linear multi-step method, for which* $p=2$, $c^*=\frac{1}{12}$. For, by (2.3), $c'=\frac{1}{12}$ implies

$$\lim_{x\to\infty} x^3r(x)/s(x)-2x^2 = 0 .$$

By (2.5) and (2.3),

$$\int_{-\infty}^{\infty} d\omega(t) = \lim_{x\to\infty} \int_{-\infty}^{\infty} \frac{x^2}{x^2+t^2} d\omega(t) = \lim_{x\to\infty} xr(x)/s(x) = 2 .$$

By this formula and (2.5),

$$\frac{x^3r(x)}{s(x)} - 2x^2 = \int_{-\infty}^{\infty} \frac{x^4}{x^2+t^2} d\omega(t) - x^2 \int_{-\infty}^{\infty} d\omega(t) = - \int_{-\infty}^{\infty} \frac{t^2x^2 d\omega(t)}{x^2+t^2} \leqq - \int_{-x}^{x} \frac{t^2x^2}{2x^2} d\omega(t) .$$

The last integral has a negative limit, unless $d\omega(t)=0$ for all $t \neq 0$. Then, by (2.4), $r(z)/s(z)=a/z$. By (2.3), $a=2$. Since $r(z)$ and $s(z)$ have no common factors, these polynomials are uniquely determined by their quotient (apart from a trivial constant factor). We already know that $r(z)/s(z)=2/z$ for the trapezoidal rule.

The concept of A-stability has an obvious meaning also outside the class of linear multistep methods. For example, the Runge–Kutta method is not A-stable, because when applied to (1.8), it gives the sequence

$$(1+qh+(qh)^2/2!+(qh)^3/3!+(qh)^4/4!)^n \qquad (n = 0,1,2,\ldots)$$

and this does not tend to zero everywhere in the half-plane $\operatorname{Re}(qh) < 0$. In fact, the base of the exponential tends to infinity, when $qh \to -\infty$. Notice, however, that *the theorems of this section are proved for linear multistep methods only*. Actually, the following modification of a linear multistep method is sufficient for the construction of an A-stable procedure of order $p = 4$.

Let $x(t,h)$ and $x(t,2h)$ be the results of the numerical integration of the same differential equation with the trapezoidal formula, using the step-size h and $2h$, respectively. Apply Richardson extrapolation *without using the extrapolated values* in the succeeding computation, i.e. for $t = 2h, 4h, 6h, 8h, \ldots$ compute

$$x^*(t,h) = x(t,h) + \tfrac{1}{3}\big(x(t,h) - x(t,2h)\big) = \tfrac{1}{3}\big(4x(t,h) - x(t,2h)\big) .$$

One can prove that, for given t, the error of $x^*(t,h)$ is $O(h^4)$, and the procedure is A-stable, since it is obtained by subtraction of the results of two A-stable procedures. Notice, however, that if the extrapolated values are used in the succeeding computation, then the A-stability is destroyed, because

$$\lim_{qh \to -\infty} \tfrac{1}{3}\left(4\left(\frac{1+\frac{1}{2}qh}{1-\frac{1}{2}qh}\right)^2 - \frac{1+qh}{1-qh}\right) = \tfrac{5}{3} > 1 .$$

3. Generalized A-stability and error estimation for the trapezoidal formula.

Consider the differential equation

$$dx/dt = f(t,x), \quad x \in R^s, \quad t_0 \leqq t < \infty \tag{3.1}$$

and make the following assumptions:

Condition A. *There exists a solution, $x = x(t)$, of class C^3 on the interval $t_0 \leqq t < \infty$.*

Condition B. *For some positive δ, the vector $f(t,x)$ and the matrix $\partial f/\partial x$ are bounded and uniformly continuous on the set,*

$$\mathfrak{S}_\delta = \{(t,x) \mid 0 < t < \infty,\ |x - x(t)| < \delta\} .$$

Introduce the modulus of continuity of $\partial f/\partial x$,

$$\omega(\varepsilon) = \sup\{|(\partial f/\partial x)_{(t,x')} - (\partial f/\partial x)_{(t,x'')}| \cdot |x' - x''| < \varepsilon,\ (t,x') \in \mathfrak{S}_\delta,\ (t,x'') \in \mathfrak{S}_\delta\} .$$

The notation $|x|$ means the *euclidean* norm of x. For the norm of a square matrix A, we write $|A| = \sup_x |Ax|/|x|$. Put

$$\boldsymbol{x}-\boldsymbol{x}(t) = \boldsymbol{y}\,,$$

$$\boldsymbol{f}\big(t,\boldsymbol{x}(t)+\boldsymbol{y}\big)-\boldsymbol{f}\big(t,\boldsymbol{x}(t)\big) = \boldsymbol{g}(t,\boldsymbol{y}) = \boldsymbol{A}(t)\boldsymbol{y}+\boldsymbol{j}(t,\boldsymbol{y})\,, \tag{3.2}$$

where

$$\boldsymbol{A}(t) = (\partial \boldsymbol{f}/\partial \boldsymbol{x})_{\boldsymbol{x}=\boldsymbol{x}(t)}\,, \tag{3.3}$$

By Condition B, $\boldsymbol{A}(t)$ is bounded, and

$$|\boldsymbol{j}(t,\boldsymbol{y})| \leqq \omega(|\boldsymbol{y}|)\cdot|\boldsymbol{y}| \leqq \omega(\delta)|\boldsymbol{y}|\,, \tag{3.4}$$

uniformly on $\mathfrak{S}_\delta$.

Next we need a generalization of the notion of A-stability. The most natural generalization would be to consider the case that $\boldsymbol{x}(t)$ is a uniform-asymptotically stable solution of (3.1) in the sense of the Liapunov theory, cf. Antosiewicz [2], but this case seems to be a little too wide. One might instead assume that the origin is uniform-asymptotically stable for the linear system

$$d\boldsymbol{y}/dt = \boldsymbol{A}(t)\boldsymbol{y} \tag{3.5}$$

where $\boldsymbol{A}(t)$ is defined by (3.3), i.e. for the so-called first approximation to (3.1) in the neighbourhood of $\boldsymbol{x}(t)$. We shall make this assumption temporarily, although it is more restrictive than necessary. It then follows from a theorem af Malkin [2], Theorem 16, $m=2$, that symmetric matrices $\boldsymbol{G}(t)$ can be found for all t, $t \geqq t_0$, such that the total derivative of the quadratic function

$$V(t,y) = \boldsymbol{y}^T\boldsymbol{G}(t)\boldsymbol{y}$$

for solutions of the linear equation (3.5) is equal to $-\boldsymbol{y}^T\boldsymbol{y}$, i.e.

$$\boldsymbol{y}^T(\boldsymbol{G}(t)\boldsymbol{A}'(t)+\boldsymbol{A}^T(t)\boldsymbol{G}(t)+\partial\boldsymbol{G}/\partial t)\boldsymbol{y} = -\boldsymbol{y}^T\cdot\boldsymbol{y} \tag{3.6}$$

$\boldsymbol{G}(t)$ and $\partial\boldsymbol{G}/\partial t$ are bounded and uniformly continuous for $t \geqq t_0$, and there exist positive constants α, β, γ, such that the following inequalities hold for all $t \geqq t_0$ and for all non-zero vectors $\boldsymbol{z}$,

$$\alpha^2|\boldsymbol{z}|^2 \leqq V(t,\boldsymbol{z}) \leqq \beta^2|\boldsymbol{z}|^2 \tag{3.7}$$

$$|\partial\boldsymbol{G}/\partial t| \leqq \alpha^2\gamma\,. \tag{3.8}$$

In other words, the function V is a quadratic Liapunov function for (3.5). It is, however, a Liapunov function for (3.1) as well. For if $\boldsymbol{y}=\boldsymbol{x}-\boldsymbol{x}(t)$, then the total derivative of V for solutions of (3.1) is equal to

$$dV/dt = \boldsymbol{y}^T\boldsymbol{G}(t)\boldsymbol{g}(t,\boldsymbol{y})+\boldsymbol{g}^T(t,\boldsymbol{y})\boldsymbol{G}(t)\boldsymbol{y}+\boldsymbol{y}^T\partial\boldsymbol{G}/\partial t\,\boldsymbol{y}\,. \tag{3.9}$$

By (3.2), (3.6) and (3.4),

$$\begin{aligned} dV/dt = {} & \boldsymbol{y}^T\boldsymbol{G}(t)\boldsymbol{A}(t)\boldsymbol{y}+\boldsymbol{y}^T\boldsymbol{A}^T(t)\boldsymbol{G}(t)\boldsymbol{y}+\boldsymbol{y}^T\partial\boldsymbol{G}/\partial t\,\boldsymbol{y}+\boldsymbol{y}^T\boldsymbol{G}(t)\boldsymbol{i}(t,\boldsymbol{y}) \\ & +\boldsymbol{j}^T(t,\boldsymbol{y})\boldsymbol{G}(t)\boldsymbol{y} = -\boldsymbol{y}^T\boldsymbol{y}+2\boldsymbol{y}^T\boldsymbol{G}(t)\boldsymbol{j}(t,\boldsymbol{y}) \leqq -|\boldsymbol{y}|^2+2\omega(\delta)|\boldsymbol{y}|^2\beta^2\,. \end{aligned}$$

Hence dV/dt is negative definite, if δ is small enough, and it follows from [2], Theorem 13, that $x(t)$ is a uniform-asymptotically stable solution of (3.1).

Now consider the following condition:

CONDITION C. *There exists a quadratic form* $V(t, y)=y^TG(t)y$, *such that, if*

$$dy/dt = g(t,y) = f(t,x(t)+y)-f(t,x(t)) ,$$

then its total derivative is negative definite on $\mathfrak{S}_\delta$. *The matrix-valued function* $G(t)$ *should satisfy* (3.7) *and* (3.8) *and* $\partial G/\partial t$ *should be uniformly continuous for* $t \geqq t_0$.

We have seen that condition C is not harder than the second of the suggestions for generalization made above. In fact, it is less restrictive. Consider for instance the scalar equation

$$dy/dt = -y^3 ,$$

and put $V(t,y)=y^2$. Then $dV/dt=-2y^4$. Condition C is obviously satisfied, although the origin is only non-asymptotically stable for the first approximation, which reads $dy/dt=0$ in this case.

Now we shall investigate the application of the trapezoidal rule to the computation of the solution $x=x(t)$. We confine ourselves to this method because of the minimum property shown in Section 2. We compute a sequence of vectors, $x_0,x_1,x_2,\ldots$ from the difference equation,

$$x_{n+1}-x_n = \tfrac{1}{2}h\big(f(t_{n+1},x_{n+1})+f(t_n,x_n)\big)+p_n ,$$

where $p_0,p_1,\ldots$ are perturbations, due for instance to roundoff errors. p_n is the error in the determination of $x_n+\frac{1}{2}hf(t_n,x_n)$, when $x_n-\frac{1}{2}hf(t_n,x_n)$ is given. The existence of x_{n+1}, when x_n and p_n are given is not clear a priori. The equation may also be written

$$x_{n+1}-\tfrac{1}{2}hf(t_{n+1},x_{n+1}) = x_n+\tfrac{1}{2}hf(t_n,x_n)+p_n . \tag{3.10}$$

Assume, however, that there exists another sequence of vectors, x_0', x_1', ..., associated with another set of perturbations, $p_0',p_1',\ldots$ satisfying the equation

$$x'_{n+1}-\tfrac{1}{2}hf(t_{n+1},x'_{n+1}) = x_n'+\tfrac{1}{2}hf(t_n,x_n')+p_n' .$$

(We shall later put $x_n'=x(t_n)$, in which case p_n' is equal to the local truncation error, but we do not specialize yet.) Put

$$f(t,z+y)-f(t,z) = g(t,y,z) . \tag{3.11}$$

Note that $g(t,y,x(t))=g(t,y)$. For $(t,z)\in\mathfrak{S}_\delta$, $(t,y+z)\in\mathfrak{S}_\delta$, we have

$$|g(t,y,z)-g(t,y)| \leqq \omega\big(|z-x(t)|\big)\,|y| \,. \tag{3.12}$$

Put, for $n=0,1,2,\ldots,$

$$y_n = x_n - x_n', \; g_n = g(t_n,y_n,x_n'), \; q_n = p_n - p_n' \,. \tag{3.13}$$

Then

$$y_{n+1} - \tfrac{1}{2}hg_{n+1} = y_n + \tfrac{1}{2}hg_n + q_n \,. \tag{3.14}$$

Define a family of norms, $|y|_n$, $n=0,1,2,\ldots,$ by

$$|y|_n{}^2 = y^T G(t_n) y \,, \tag{3.15}$$

where, for each n, $G(t_n)$ is a positive definite, symmetric matrix. (We shall later impose further conditions on these matrices.) By the triangular inequality,

$$|y_{n+1} - \tfrac{1}{2}hg_{n+1}|_n \leqq |y_n + \tfrac{1}{2}hg_n|_n + |q_n|_n \,. \tag{3.16}$$

Now, let y,z be two arbitrary vectors, and put $g=g(t,y,z)$. Then

$$\begin{aligned} &|y+\tfrac{1}{2}hg|_n{}^2 - |y-\tfrac{1}{2}hg|_{n-1}^2 \\ &\quad = |y+\tfrac{1}{2}hg|_n{}^2 - |y-\tfrac{1}{2}hg|_n{}^2 + |y-\tfrac{1}{2}hg|_n{}^2 - |y-\tfrac{1}{2}hg|_{n-1}^2 \\ &\quad = hW(t_n,y,z,h) \end{aligned} \tag{3.17}$$

where W is defined by

$$W(t,y,z,h) = 2y^T G(t) g + (y-\tfrac{1}{2}hg)^T \frac{G(t)-G(t-h)}{h} (y-\tfrac{1}{2}hg) \,, \tag{3.18}$$

Note that, by (3.9),

$$\lim_{h\to 0} W\big(t,y,x(t),h\big) = dV/dt \,.$$

The last relations give a motivation for the following modification of Condition C.

Condition C_λ'. *Given h, δ, λ. There should exist a symmetric matrix $G(t)$ satisfying* (3.7) *and* (3.8), *for all t, $t \geqq t_0$ such that for all $(t,z) \in \mathfrak{S}_\delta$, $(t,y+z) \in \mathfrak{S}_\delta$, $0<u\leqq h$,*

$$W(t,y,z,u) \leqq 2\lambda(|y|^2 + |\tfrac{1}{2}ug|^2) \,. \tag{3.19}$$

In many cases, the same $G(t)$ can be used for the differential and the difference equation. For example: if the differential equation is linear, and if there exists a time-independent, quadratic Liapunov function, $V=y^TGy$, then, by (3.9) and (3.11), $W(t,y,z,u)=dV/dt$. It can then be shown that there exists a negative λ, such that C_λ' is true for all h,δ. If there exists a Liapunov function of the same kind for the first approximation (3.5) to a non-linear system (3.1) then there exists a negative λ, such that C_λ' is true for all h and all sufficiently small δ.

We shall always assume that A, B, C_λ' are true for the values of h, δ, λ under consideration, and that $(t,\boldsymbol{z}) \in \mathfrak{S}_\delta$, $(t,\boldsymbol{y}+\boldsymbol{z}) \in \mathfrak{S}_\delta$. We need a few notations, identities and inequalities. Consider the well-known identity

$$2(|\boldsymbol{y}|^2+|\tfrac{1}{2}h\boldsymbol{g}|^2) = |\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|^2+|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|^2 . \tag{3.20}$$

Hence, by (3.5) and (3.15),

$$\begin{aligned}\beta^{-2}(|\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|_n{}^2+|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1}^2) &\leqq 2(|\boldsymbol{y}|^2\\ &+|\tfrac{1}{2}h\boldsymbol{g}|^2) \leqq \alpha^{-2}(|\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|_n{}^2+|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1}^2) .\end{aligned} \tag{3.21}$$

Put

$$\lambda' = \begin{cases}\lambda\alpha^{-2}, \text{ if } \lambda \geqq 0 ,\\ \lambda\beta^{-2}, \text{ if } \lambda \leqq 0 .\end{cases}$$

By (3.17), (3.19) and (3.21), if $\boldsymbol{g}=\boldsymbol{g}(t_n,\boldsymbol{y},\boldsymbol{z})$,

$$|\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|_n{}^2-|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1}^2 = hW(t_n,\boldsymbol{y},\boldsymbol{z},h) \leqq h\lambda'(|\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|_n{}^2+|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1}^2) . \tag{3.22}$$

Hence

$$|\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|_n{}^2 \leqq \frac{1+h\lambda'}{1-h\lambda'}\,|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1}^2, \text{ if } h\lambda' < 1 . \tag{3.23}$$

By (3.21) and (3.23),

$$\begin{aligned}2|\boldsymbol{y}|^2 \leqq 2(|\boldsymbol{y}|^2+|\tfrac{1}{2}h\boldsymbol{g}|^2) &\leqq \alpha^{-2}\big((1+h\lambda')/(1-h\lambda')+1\big)|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1}^2\\ |\boldsymbol{y}| &\leqq (1-h\lambda')^{-\frac{1}{2}}\alpha^{-1}|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1}, \text{ if } h\lambda' < 1 .\end{aligned} \tag{3.24}$$

Now put

$$\mu = \frac{1}{2h}\log\frac{1+h\lambda'}{1-h\lambda'}, \text{ if } h\lambda' < 1 . \tag{3.25}$$

By (3.23),

$$|\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|_n \leqq e^{\mu h}|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1} , \tag{3.26}$$

and hence, by (3.16), we obtain the important inequality

$$|\boldsymbol{y}_{n+1}-\tfrac{1}{2}h\boldsymbol{g}_{n+1}|_n \leqq e^{\mu h}|\boldsymbol{y}_n-\tfrac{1}{2}h\boldsymbol{g}_n|_{n-1}+|\boldsymbol{q}_n|_n . \tag{3.27}$$

Let q^* be an upper bound for $|\boldsymbol{p}_n|+|\boldsymbol{p}_n'|$. Note that

$$|\boldsymbol{q}_n|_n \leqq \beta|\boldsymbol{q}_n| \leqq \beta q^* . \tag{3.28}$$

Let ε_0 be an upper bound of the errors in $\boldsymbol{x}_0(1-h\lambda')^{\frac{1}{2}}$ and $\boldsymbol{x}_0+\frac{1}{2}h\boldsymbol{f}(t_0,\boldsymbol{x}_0)$, and put

$$\Phi(t,h) = \varepsilon_0 e^{\mu(t-h)}+q^*\cdot\frac{1-e^{\mu t}}{1-e^{\mu h}} . \tag{3.29}$$

It is easily verified that $\Phi_n=\Phi(t_n,h)$ is the solution of the difference equation

$$\Phi_{n+1} = e^{\mu h}\Phi_n + q^*, \quad \Phi_1 = \varepsilon_0 + q^* . \tag{3.30}$$

Comparing (3.27) and (3.30), we find by induction that

$$|\boldsymbol{y}_n - \tfrac{1}{2}h\boldsymbol{g}_n|_{n-1} \leqq \beta\Phi(t_n,h), \quad (n \geqq 1) \tag{3.31}$$

(which leads to a bound for $|\boldsymbol{y}_n|$ by use of (3.24)), provided that we can be sure that, at each step, (3.10) has a solution $(t_{n+1},\boldsymbol{x}_{n+1}) \in \mathfrak{S}_\lambda$. However, we have to worry about this, because the iterative method, which is usually applied in a constructive existence proof, converges, roughly speaking, only if all eigenvalues of $\frac{1}{2}h\partial\boldsymbol{f}/\partial\boldsymbol{x}$ are located inside the unit circle, which is an unsatisfactory restriction. By means of a modification (under-relaxation), this procedure may be extended to the case where the real part of every eigenvalue of $\frac{1}{2}h\partial\boldsymbol{f}/\partial\boldsymbol{x}$ is less than unity. This might be applied here, at least to the case with a constant Liapunov function. We shall, however, proceed in a different way.

Lemma 3.1. *Given $\boldsymbol{x}_n$, $\boldsymbol{p}_n$. If $h\lambda' < 1$ (which means no restriction on h, if $\lambda \leqq 0$), then the equation* (3.10) *has at most one solution, such that* $(t_{n+1},\boldsymbol{x}_{n+1}) \in \mathfrak{S}_\delta$.

Proof. If there were two different solutions, $\boldsymbol{x}_{n+1}$, $\boldsymbol{x}'_{n+1}$, with $\boldsymbol{p}_n = \boldsymbol{p}'_n$, then $\boldsymbol{y} = \boldsymbol{x}_{n+1} - \boldsymbol{x}'_{n+1}$ would satisfy

$$\boldsymbol{y} - \tfrac{1}{2}h\boldsymbol{g}(t_{n+1},\boldsymbol{y},\boldsymbol{x}'_{n+1}) = \boldsymbol{0} .$$

We find that $\boldsymbol{y} = \boldsymbol{0}$, by substituting $n+1$ for n in (3.24).

Lemma 3.2. *If $h\lambda' < 1$, then the matrix $\boldsymbol{I} - \frac{1}{2}h\partial\boldsymbol{f}/\partial\boldsymbol{x}$ is non-singular, for all points $(t,\boldsymbol{x}) \in \mathfrak{S}_\delta$.*

Proof. If it were singular, there would exist vectors $\boldsymbol{y}$ of any length such that $\boldsymbol{y} - \frac{1}{2}h\partial\boldsymbol{f}(t,\boldsymbol{x})/\partial\boldsymbol{x}\cdot\boldsymbol{y} = 0$. Hence, by (3.11), for any ε, there would exist a vector $\boldsymbol{y}$ such that

$$|\boldsymbol{y} - \tfrac{1}{2}h\boldsymbol{g}(t,\boldsymbol{y},\boldsymbol{x})| < \varepsilon|\boldsymbol{y}| ,$$

but this is impossible, according to (3.24).

Now, put $\boldsymbol{x}_n' = \boldsymbol{x}(t_n)$. Then, $\boldsymbol{p}_n'$ is the local truncation error of the trapezoidal rule, which is known to be less than $h^3 \sup|\boldsymbol{x}'''(t)|/12$. Hence we may put

$$q^* = \text{u.b.}|\boldsymbol{p}_n| + h^3 \sup|\boldsymbol{x}'''(t)|/12 . \tag{3.32}$$

Lemma 3.3. *Given h, δ, λ, n, $(t_n,\boldsymbol{x}_n) \in \mathfrak{S}_\delta$, $\boldsymbol{p}_n$, $\boldsymbol{x}_n' = \boldsymbol{x}(t_n)$, $\varepsilon > 0$. Suppose that $h\lambda' < 1$ and that*

$$|\boldsymbol{y}_n|_n \leqq \alpha\delta - \beta q^* - \varepsilon, \quad |\boldsymbol{y}_n + \tfrac{1}{2}h\boldsymbol{g}(t_n,\boldsymbol{y}_n)|_n \leqq \alpha\delta(1-h\lambda')^{\frac{1}{2}} - \beta q^* - \varepsilon . \tag{3.33}$$

Then (3.10) *has a solution,* $(t_{n+1},\boldsymbol{x}_{n+1}) \in \mathfrak{S}_\delta$.

PROOF. For given t_n, substitute u for h in (3.10) and the equivalent equation (3.14), which then reads

$$\boldsymbol{y}_{n+1}-\tfrac{1}{2}u\boldsymbol{g}_{n+1} = \boldsymbol{y}_n+\tfrac{1}{2}u\boldsymbol{g}_n+\boldsymbol{q}_n,\ \boldsymbol{g}_{n+1} = \boldsymbol{g}(t_n+u,\boldsymbol{y}_{n+1})\,. \qquad (3.33')$$

Assume that $0 \leqq u \leqq h$, $t_{n+1}=t_n+u$, $\boldsymbol{x}'_{n+1}=\boldsymbol{x}(t_n+u)$, $\boldsymbol{q}_n=\boldsymbol{p}_n-\boldsymbol{p}_n'$ depends on u, but by (3.32) it is still true that $|\boldsymbol{q}_n| \leqq q^*$, if h is replaced by a smaller quantity. The solutions of (3.10), (3.33) will be considered as functions $\boldsymbol{x}_{n+1}(u)$, $\boldsymbol{y}_{n+1}(u)$. Also note that C_λ' holds, when h is replaced by u. Hence, we may substitute u for h in all the results of local type we have derived so far.

Clearly,

$$|\boldsymbol{y}_{n+1}(0)| = |\boldsymbol{y}_n+\boldsymbol{q}_n| \leqq \alpha^{-1}|\boldsymbol{y}_n|_n+q^* < \delta-\beta q^*/\alpha+q^* \leqq \delta\,,$$

by (3.33). Hence $\big(t_n,\boldsymbol{x}_{n+1}(0)\big) \in \mathfrak{S}_\delta$. When u increases, $\boldsymbol{x}_{n+1}(u)$ is continuous, by Lemma 3.2 and the existence theorem for implicit functions, as long as $\big(t_n+u,\boldsymbol{x}_{n+1}(u)\big)$ stays in $\mathfrak{S}_\delta$, i.e. as long as $|\boldsymbol{y}_{n+1}(u)|<\delta$. Now, consider the inequalities (3.33). We can interpolate between them, because $|\boldsymbol{y}_n+\frac{1}{2}u\boldsymbol{g}_n|_n$ is a convex function of u, while $(1-u\lambda')^{\frac{1}{2}}$ is concave. Hence

$$|\boldsymbol{y}_n+\tfrac{1}{2}u\boldsymbol{g}_n|_n \leqq \alpha\delta(1-u\lambda')^{\frac{1}{2}}-\beta q^*-\varepsilon\,.$$

It follows from this, (3.33′) and (3.28), that

$$|\boldsymbol{y}_{n+1}(u)-\tfrac{1}{2}u\boldsymbol{g}_{n+1}|_n \leqq \alpha\delta(1-u\lambda')^{\frac{1}{2}}-\varepsilon$$

and, by (3.24),

$$|\boldsymbol{y}_{n+1}(u)| \leqq \delta-\alpha^{-1}(1-u\lambda')^{-\frac{1}{2}}\varepsilon\,.$$

This shows that $\big(t_n+u,\boldsymbol{x}_{n+1}(u)\big)$ will not be able to reach the boundary of $\mathfrak{S}_\delta$, under the assumptions made. Hence (3.10) has a solution in $\mathfrak{S}_\delta$.

We shall now obtain sufficient conditions for the validity of (3.31), but first we need one more inequality. Let $\boldsymbol{x}$ be an arbitrary vector. Then

$$|\boldsymbol{x}|_n^2-|\boldsymbol{x}|_{n-1}^2 = \boldsymbol{x}^T\big(G(t_n)-G(t_n-h)\big)\boldsymbol{x} = -\boldsymbol{x}^T\int_0^h \frac{\partial G(t_n-\tau)}{\partial\tau}\,d\tau\,\boldsymbol{x}$$
$$\leqq |\boldsymbol{x}|^2h\alpha^2\gamma \leqq h\gamma|\boldsymbol{x}|_{n-1}^2\,,$$

by (3.6) and (3.5). Hence

$$|\boldsymbol{x}|_n{}^2 \leqq (1+h\gamma)|\boldsymbol{x}|_{n-1}^2\,. \qquad (3.34)$$

THEOREM 3.1. *Given h, δ, λ. Assume that*

(i) $h\lambda'<1$, (*which means no restriction if* $\lambda \leqq 0$),

(ii) C_λ' *is satisfied,*

(iii) *for all* t, $t_0 \leqq t < T \leqq \infty$,

$$\Phi(t,h) < (\alpha\delta/\beta - q^*)\big((1-h\lambda')^{-1}+\tfrac{1}{2}h\gamma\big)^{-\frac{1}{2}} .$$

Then, (3.10) *uniquely defines a sequence* $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots \boldsymbol{x}_n,\ldots$ *as long as* $t_0+nh \leqq T$. *The following bound is valid:*

$$|\boldsymbol{x}_n - \boldsymbol{x}(t_n)| \leqq \Phi(t_n,h)(\beta/\alpha)(1-h\lambda')^{-\frac{1}{2}} . \tag{3.35}$$

Remark. Expressions for $\Phi(t,h)$ and q^* are found in (3.29) and (3.32).

Proof. Assume that $\boldsymbol{y}_n$ exists and that $|\boldsymbol{y}_n| < \delta$, for $n=1,2,\ldots,m$. Hence (3.31) holds for $n \leqq m$,

$$|\boldsymbol{y}_n - \tfrac{1}{2}h\boldsymbol{g}_n|_{n-1} \leqq \beta\Phi(t_n,h) \tag{3.36}$$

and (3.35) is obtained by an application of (3.24). We now only need to show that the conditions for the existence of a point $(t_{m+1},\boldsymbol{x}_{m+1}) \in \mathfrak{S}_\delta$, given in Lemma 3.3, are satisfied. (The verification for $m=1$ is straightforward, by (3.29) and (iii)). By (3.26), (3.36), (3.30) and (iii),

$$|\boldsymbol{y}_m + \tfrac{1}{2}h\boldsymbol{g}_m|_m \leqq e^{\mu h}\beta\Phi(t_m,h) = \beta\Phi(t_{m+1},h) - q^*\beta < \alpha\delta(1-h\lambda')^{\frac{1}{2}} - \beta q^* . \tag{3.37}$$

Consider the identity

$$2(|\boldsymbol{y}_m|_m{}^2 + |\tfrac{1}{2}h\boldsymbol{g}_m|_m{}^2) = |\boldsymbol{y}_m + \tfrac{1}{2}h\boldsymbol{g}_m|_m{}^2 + |\boldsymbol{y}_m - \tfrac{1}{2}h\boldsymbol{g}_m|_m{}^2 .$$

By (3.23), (3.34) and (3.36),

$$2|\boldsymbol{y}_m|_m{}^2 \leqq \left(\frac{1+h\lambda'}{1-h\lambda'} + 1 + h\gamma\right)\beta^2\Phi^2(t_m,h) .$$

Hence

$$|\boldsymbol{y}_m|_m \leqq \big((1-h\lambda')^{-1}+\tfrac{1}{2}h\gamma\big)^{\frac{1}{2}}\beta\Phi(t_m,h) < (\alpha\delta - \beta q^*) ,$$

by (iii). This inequality and (3.37) are equivalent to the conditions of Lemma 3.3, since they are strict inequalities. Hence (3.10) has a solution $(t_{m+1},\boldsymbol{x}_{m+1}) \in \mathfrak{S}_\delta$, which is unique, by Lemma 3.1. Hence (3.31) holds also for $n=m+1$, and the theorem is proved.

The classical error estimates contained exponentials with a Lipschitz constant, essentially $|\partial \boldsymbol{f}/\partial \boldsymbol{x}|$, as the coefficient of t in the exponent. In recent years, several writers have derived error bounds, where the exponents may be negative, just like μ in our error bound, when $\lambda < 0$. Some bounds of this kind are found in [3, Ch. 5], but, in connection with the trapezoidal formula, they are less general than Theorem 3.1. For example, they do not always give sharp bounds, when Liapunov functions with variable coefficients are needed, and, above all, they are based on the assumption $\frac{1}{2}h|\partial \boldsymbol{f}/\partial \boldsymbol{x}| < 1$. According to Theorem 3.1, when $\lambda < 0$ the choice of step size can be made with consideration of the local truncation error and the rate of change of the Liapunov function only. Even though $|h\partial \boldsymbol{f}/\partial \boldsymbol{x}|$ is large, the error is bounded in an unlimited computation, if $\lambda < 0$. *This is a generalization of the A-stability property.*

If $\lambda = 0$, $\Phi(t,h)$ grows linearly with t, because

$$\lim_{\lambda \to 0} \frac{1-e^{\mu t}}{1-e^{\mu h}} = \frac{t}{h}\,.$$

There are, however, important cases, where the error is bounded, when $\lambda = 0$, although Theorem 3.1 fails to indicate this. Let us introduce a new condition.

CONDITION C_0''. *In addition to condition C_0', it is required that, for $|\boldsymbol{y}| < \delta$, $W(t,\boldsymbol{y},\boldsymbol{x}(t),h)$ is not larger than some continuous, negative definite (though in general not quadratic) time-independent function of the vector $\boldsymbol{y} - \frac{1}{2}h\boldsymbol{g}$, say*

$$W(t,\boldsymbol{y},\boldsymbol{x}(t),h) \leqq -2|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1} W_1(|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1})\,,$$

where $\boldsymbol{g} = \boldsymbol{g}(t,\boldsymbol{y})$, $W_1(s) > 0$ for $s \neq 0$.

We may use any result, that has been obtained under Condition C_0'. Hence, by (3.22),

$$|\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|_n^2 - |\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1}^2 \leqq -2h|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1} W_1(|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1})\,.$$

Divide by $|\boldsymbol{y}+\frac{1}{2}h\boldsymbol{g}|_n + |\boldsymbol{y}-\frac{1}{2}h\boldsymbol{g}|_{n-1} \leqq 2|\boldsymbol{y}-\frac{1}{2}h\boldsymbol{g}|_{n-1}$. Hence

$$|\boldsymbol{y}+\tfrac{1}{2}h\boldsymbol{g}|_n - |\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1} \leqq -hW_1(|\boldsymbol{y}-\tfrac{1}{2}h\boldsymbol{g}|_{n-1})\,.$$

(It is now seen that $W_1(0) = 0$.) *Put*

$$|\boldsymbol{y}_n - \tfrac{1}{2}h\boldsymbol{g}_n|_{n-1} = s_n\,.$$

It now follows from (3.14) and (3.28) that

$$s_{n+1} \leqq s_n - hW_1(s_n) + \beta q^*\,. \tag{3.38}$$

Now we are going to prove:

THEOREM 3.2. *Given h, δ. Assume that:*

(i) C_0'' *is satisfied,*

(ii) $\beta q^*/h < \underline{\lim}\, W_1(s)$.

Put $M = \sup\{s - hW_1(s) + \beta q^* \mid 0 \leqq s \leqq s'\}$
where s' is the largest root of the equation $hW_1(s) = \beta q^*$.

(iii) $M < (\alpha\delta - \beta q^*)(1+\frac{1}{2}h\gamma)^{-\frac{1}{2}}$,

(iv) $s_1 \leqq M$.

Then $s_n \leqq M$ for all n, and

$$|x_n - x(t_n)| \leqq \alpha^{-1}M < \delta\,. \tag{3.39}$$

Proof. Note that $s' \leqq M$, and that $hW_1(s) > \beta q^*$, for $s > s'$. If $s_n \leqq s'$, then $s_{n+1} \leqq s_n - hW_1(s_n) + \beta q^* \leqq M$. If $s_n > s'$, then $s_{n+1} \leqq s_n + (-hW_1(s_n) + \beta q^*) < s_n$. Hence, in any case, if $s_n \leqq M$, then $s_{n+1} \leqq M$. All this is based on the assumption that, at each step, (3.10) has a solution in $\mathfrak{S}_\delta$. The proof of this assumption is obtained by the substitution of M for $\beta\Phi(t,h)$ in the proof of Theorem 3.1. (3.39) is obtained by an application of (3.24).

Theorem 3.3. *Make the same assumptions as in the preceding theorem, except that* (iv) *is replaced by the milder assumption*

(iv′) $s_1 < (\alpha\delta - \beta q^*)(1 + \tfrac{1}{2}h\gamma)^{-\frac{1}{2}}$.

Then, as $n \to \infty$, $\overline{\lim}\, s_n \leqq M$ *and* $\overline{\lim}\, |x_n - x(t_n)| \leqq \alpha^{-1}M$.

Proof. If, for some n, $s_n \leqq M$, then the statement follows from the preceding theorem. Therefore, assume that $s_n > M$ for all n. Then, a fortiori, $s_n > s'$, so that $s_{n+1} \leqq s_n - hW_1(s_n) + \beta q^* < s_n$. Hence $\{s_n\}$ is a bounded, decreasing sequence, which tends to some limit, s'' satisfying the inequality $s'' \leqq s'' - hW_1(s'') + \beta q^*$. Hence $hW_1(s'') \leqq \beta q^*$, whence $s'' \leqq s' \leqq M$, and the theorem is proved. The argument also shows that, *if* $M \neq s'$ *then* $s_n \leqq M$ *for all sufficiently large* n.

A simple example, where C_0'' is satisfied although C_λ' is not, for any negative λ, is given by the equation $dy/dt = -y^3$. Take $V(t) = y^2$. Then, $W(t,y,z) = -2y\big((y+z)^3 - z^3\big) = -2y^2(y^2 + yz + z^2) \leqq -y^2(y^2 + z^2) \leqq -y^4$. On the other hand, $W(t,y,0) = -2y^4 \geqq -2\delta^2 y^2$, when $|y| < \delta$.

The remarkable stability property of the trapezoidal formula has to be matched against two obvious disadvantages:

1. p is only equal to 2,
2. the method is implicit.

To some extent, one can compensate for the first disadvantage by use of Richardson extrapolation. In many cases, the second difficulty may be overcome by a suitable combination of elimination and iteration, although in other cases it may be more economical to use an explicit method and a smaller step-size. The previous theory is applicable, when the computations are arranged so that the sum of the perturbation $|p_n|$ in (3.10) and the local truncation error $|p_n'|$ never exceeds some fixed bound called q^*. It is, however, important to realize that any condition of this type may be violated, eventually, if an iterative technique is used with a *fixed* number of iterations, and the error $|x_n - x(t_n)|$ may grow to infinity, even though $|h\partial f/\partial x|$ is rather small.

It is instructive to study the application of the iterative scheme

$$\begin{aligned} x_{n+1}^{(0)} &= x_n \\ x_{n+1}^{(i)} &= x_n + \tfrac{1}{2}h\big(f(x_n) + f(x_{n+1}^{(i-1)})\big), \qquad i = 1,2,\ldots j: \end{aligned}$$

for different constant values of j, in the case $f(x) = qx$, where q is located on or very close to the imaginary axis. The boundedness of the sequence $\{x_n^{(j)}\}_{n=1}^{\infty}$ for fixed j and fixed imaginary values of qh of small modulus, then depends on the residue of j modulo 4. Prof. Herbert Keller, New York, pointed out this peculiar fact to the writer.

Acknowledgement.

The writer had the privilege to prepare part of this paper as a visitor at the University of California, Los Angeles, and the Courant Institute of Mathematical Sciences, New York. I am much indebted to Professors Peter Henrici and Peter Lax for stimulating discussions.

Note, added in proof.

It ought to be mentioned that *if an A-stable method is applied to* (1.8) *with purely imaginary q, then all solutions are bounded.* By continuity, the roots of (2.1) satisfy the condition $|\zeta| \le 1$, and by the technique used in the proof of Lemma 2.1, it can be shown that the roots of unit modulus are simple. (Note that $\varrho(\zeta)/\sigma(\zeta) - qh$ has a non-negative real part for $|\zeta| > 1$.)

REFERENCES

1. Achieser, N. I. — Glassman, I. M., *Theorie der linearen Operatoren im Hilbert-Raum,* Akademie-Verlag, Berlin 1954.
2. Antosiewicz, H.-A., *A survey of Liapunov's second method,* in Contributions to the theory of non-linear oscillation, vol. IV., S. Lefschetz (ed.), Ann. Math. Studies, No. 41, Ch VIII, Princeton 1958.
3. Dahlquist, G., *Stability and error bounds in the numerical integration of ordinary differential equations,* Dissertation, Stockholm 1958. Also in Trans. Roy. Inst. Technol. Stockholm, Nr. 130, (1959).
4. Dahlquist, G., *Stability questions for some numerical methods for ordinary differential equations,* To appear in Proc. Symposia on Applied Mathematics, vol. 15, „Interactions between Mathematical Research and High-Speed Computing, 1962."
5. Hamming, R. W., *Numerical methods for scientists and engineers,* McGraw-Hill, 1962.
6. Henrici, P. K., *Discrete variable methods in ordinary differential equations,* Wiley, 1962.
7. Robertson, H. H., *Some new formulae for the numerical integration of ordinary differential equations,* Information Processing, UNESCO, Paris, pp. 106–108.

ROYAL INSTITUTE OF TECHNOLOGY,
STOCKHOLM, SWEDEN